What is claimed is:

1. A method for determining a gesture comprising the steps of:

determining a change in a background of an image from a plurality of images;

determining an object in the image;

determining a trajectory of the object through the plurality of images; and

classifying a gesture according to the trajectory of the object.

2. The method of claim 1, wherein the step of determining the change in the background further comprises the steps of:

determining a gradient intensity map for the background from a plurality of images;

determining a gradient intensity map for the current image;

determining, for a plurality of pixels, a difference between the gradient intensity map and the gradient intensity map for the background;

determining a comparison between the difference and a threshold; and

determining a pixel to be a background pixel according to the comparison.

3. The method of claim 1, wherein the object includes a user's hand.

4. The method of claim 1, wherein the step of determining the object in the image further comprises the steps of:

obtaining a normalized color representation for a plurality of colors in each image;

determining from training images an estimate of a probability distribution of normalized color values for an object class; and

determining, for each pixel, a likelihood according to an estimated probability density of normalized color values for the object class.

5. The method of claim 1, wherein the step of determining the trajectory of the object through the plurality of images further comprises the steps of:

determining, for each pixel, a temporal likelihood across a plurality of images; and

determining a plurality of moments according to the temporal likelihoods.

6. The method of claim 1, wherein the step of determining the trajectory further comprises the steps of:

determining a difference in a size of the object over a pre-determined time period;

determining a plurality of angles between a plurality of lines connecting successive controids over the time period; and

determining a feature vector according to the angles and lines.

7. The method of claim 6, further comprising the step of classifying the feature vector according to a time-delay neural network, wherein a feature is of a fixed length.

8. The method of claim 1, wherein the step of classifying the gesture further comprises the steps of:

determining a reference point;

determining a correspondence between the trajectory and the reference point; and

classifying the trajectory according to one of a plurality of commands.

9. A method for determining a trajectory of a hand through a plurality of images comprising the steps of:

detecting a reference point;

updating the reference point as the reference point is varied;

detecting a first translation of the hand through the plurality of images;

detecting a second translation through the plurality of images;

determining a gesture according a vote; and

determining whether the gesture is a valid gesture command.

10. The method of claim 9, wherein the reference point is not interpreted as a gesture command.

11. The method of claim 9, wherein the reference point is characterized by hand size and a location of a centroid of the hand in each image.

12. The method of claim 9, wherein the first translation is one of a forward and a backward translation, wherein the first translation is characterized by a large change in hand size and a relatively small change in a centroid of the hand.

13. The method of claim 9, wherein the second translation is one of a left, a right, an up and a down translation.

14. The method of claim 9, wherein the step of detecting the second translation further comprises the step of determining a normalized vector between two centroids c_t and c_{t-1} as a feature vector, wherein there are three output patterns.

15. The method of claim 14, wherein the three output patterns are a vertical movement, a horizontal movement, and an unknown, the method further comprising the steps of:

comparing the reference point to a centroid upon determining the translation to be a vertical or a horizontal translation; and

testing an input pattern upon determining the translation to be an unknown translation.

16. The method of claim 15, wherein the step of testing an input pattern further comprises the steps of detecting a circular movement, wherein an angle between vector $c_t c_{t-1}$ and vector $c_{t-1} c_{t-2}$ is determined as the feature vector.

17. The method of claim 9, wherein the valid gesture is performed continually for a predetermined time.

18. A program storage device readable by machine, tangibly embodying a program of instructions executable by the machine to perform method steps for determining a gesture, the method steps comprising:

determining a change in a background of an image from a plurality of images;

determining an object in the image;

determining a trajectory of the object through the plurality of images; and

classifying a gesture according to the trajectory of the object.

19. The method of claim 18, wherein the step of determining the change in the background further comprises the steps of:

determining a gradient intensity map for the background from a plurality of images;

determining a gradient intensity map for the current image;

determining, for a plurality of pixels, a difference between the gradient intensity map and the gradient intensity map for the background;

determining a comparison between the difference and a threshold; and

determining a pixel to be a background pixel according to the comparison.

20. The method of claim 18, wherein the object includes a user's hand.

21. The method of claim 18, wherein the step of determining the object in the image further comprises the steps of:

obtaining a normalized color representation for a plurality of colors in each image;

determining from training images an estimate of a probability distribution of normalized color values for an object class; and

determining, for each pixel, a likelihood according to an estimated probability density of normalized color values for the object class.

22. The method of claim 18, wherein the step of determining the trajectory of the object through the plurality of images further comprises the steps of:

determining, for each pixel, a temporal likelihood across a plurality of images; and

determining a plurality of moments according to the temporal likelihoods.

23. The method of claim 18, wherein the step of determining the trajectory further comprises the steps of:

determining a difference in a size of the object over a pre-determined time period;

determining a plurality of angles between a plurality of lines connecting successive controids over the time period; and

determining a feature vector according to the angles and lines.

24. The method of claim 23, further comprising the step of classifying the feature vector according to a time-delay neural network, wherein a feature is of a fixed length.

25. The method of claim 18, wherein the step of classifying the gesture further comprises the steps of:

determining a reference point;

determining a correspondence between the trajectory and the reference point; and

classifying the trajectory according to one of a plurality of commands.